

MERCK PATENT GmbH (DE)
 Location/Qualifiers
 1. .1305
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 1. .>1305
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28479.1"
 /db_xref="GI:12710816"
 /db_xref="EMBL:CAC28479"
 /translation="MDPSDQPLNSDVPLKPKPIMETPKKVGIPPIIALSLASI
 LLSVLLKVLIDKYSFQGLPLHPIPKQCDGELDCPLGDEDEHCVKSPFEGPAV
 RLVSTLQVLDSATQWESACFDPTALAEACQMGYSKPTFAVEIGDQD
 DVETJENSDELKRNSSGCLSGSLVSLHCLACQSLKPTPRVVGEEASVDSFPMQV
 SIQDKHVCQGSILDPHWLTAHGRKHTDFNNKVRAGSKLGSFSLAVAKII
 IEFNPMYKNDIALMKLQPLPFTSGTVRPICLPFFDEBELPATPLNTIGMFTKQNG
 GKMSDILLQASVQVLDSTRNADDAVQGEVTEKMWAGIPEGVJTCQDSSGPLYQ
 SDQHWVGVISWGYCGSPSTPGVYTKVSAYLNIWVYKAEI"
 ORIGIN
 Query Match 100.0%; Score 1305; DB 6; Length 1305;
 Best Local Similarity 100.0%; Pred. No. 7, 5e-302;
 Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGATCTGACAGTATCACTCTGAACAGCTCGATGTCAACACCCCTGGCGAACCC 60
 DB 1 ATGGATCTGACAGTATCACTCTGAACAGCTCGATGTCAACACCCCTGGCGAACCC 60
 QY 61 CGTATCCCCATGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGC 120
 DB 61 CGTATCCCCATGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGC 120
 QY 121 CTGGCGAGTATCATCATTTGTGTTGCTCTCATCAAGGTGATTCGTGATTAATACTACTTC 180
 DB 121 CTGGCGAGTATCATCATTTGTGTTGCTCTCATCAAGGTGATTCGTGATTAATACTACTTC 180
 QY 181 CTCTGGGGAGCTCTCCATCTCCAGGAAGCAGCTGTGTGAGGAGCTGGAC 240
 DB 181 CTCTGGGGAGCTCTCCATCTCCAGGAAGCAGCTGTGTGAGGAGCTGGAC 240
 QY 241 TGTCCCTTGGGGAGGAGCAGGAGCCTGTGTGAGAGCTTCCCGAAGGGCTCGAGTG 300
 DB 241 TGTCCCTTGGGGAGGAGCAGGAGCCTGTGTGAGAGCTTCCCGAAGGGCTCGAGTG 300
 QY 301 CGAGTCCGCTCTCAAGGACCGATCCACATGCAAGTCTGGAATCGGGCCACACGGGAAC 360
 DB 301 CGAGTCCGCTCTCAAGGACCGATCCACATGCAAGTCTGGAATCGGGCCACACGGGAAC 360
 QY 361 TGGTCTCTGCTGCTGTTTCGACAACTTCACAGAACTCTGCTGACACAGCTGTAGGCAG 420
 DB 361 TGGTCTCTGCTGCTGTTTCGACAACTTCACAGAACTCTGCTGACACAGCTGTAGGCAG 420
 QY 421 ATGGGCTACACAGCAACCCCTTTCAGAGCTGGAGATTCGCCAGACAGATCTG 480
 DB 421 ATGGGCTACACAGCAACCCCTTTCAGAGCTGGAGATTCGCCAGACAGATCTG 480
 QY 481 GATGTTGTGTAATTCACAGAAAACAGCAGGAGCTTTCGATCGGAACTCAAGTGGGCC 540
 DB 481 GATGTTGTGTAATTCACAGAAAACAGCAGGAGCTTTCGATCGGAACTCAAGTGGGCC 540
 QY 541 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGGCTGTGGGAAGAGCTCAAGACC 600
 DB 541 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGGCTGTGGGAAGAGCTCAAGACC 600
 QY 601 CCCCGTGTGGTGGGAGGAGGCTCTGTGGATTCCTGGCTTGGCAGTCAGATC 660
 DB 601 CCCCGTGTGGTGGGAGGAGGCTCTGTGGATTCCTGGCTTGGCAGTCAGATC 660
 QY 661 CAGTACGACAGCAGCTCTGTGGAGGAGCAGCTCTGGACCCCTCGCTCTCAG 720
 DB 661 CAGTACGACAGCAGCAGCTCTGTGGAGGAGCAGCTCTGGACCCCTCGCTCTCAG 720

QY 721 GCAGCCCACTGCTTTCAGGAAACATACCGATGTTGTTCAACTGGAAGGTGCGGCGAGGCTCA 780
 DB 721 GCAGCCCACTGCTTTCAGGAAACATACCGATGTTGTTCAACTGGAAGGTGCGGCGAGGCTCA 780
 QY 781 GACAAACTGGGAGCTTCCCATCCCTGGCTGTGGCCAGAGATCATCATCTGAATTCAC 840
 DB 781 GACAAACTGGGAGCTTCCCATCCCTGGCTGTGGCCAGAGATCATCATCTGAATTCAC 840
 QY 841 CCCATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAATCTCCCACTCATTTTC 900
 DB 841 CCCATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAATCTCCCACTCATTTTC 900
 QY 901 TCAGGCACAGTCAAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGGCCACC 960
 DB 901 TCAGGCACAGTCAAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGGCCACC 960
 QY 961 CCATCTGGATCATTTGGATGGGCTTTTACGAAGCAGATGCGAGGAAGATGTTCTGACATA 1020
 DB 961 CCATCTGGATCATTTGGATGGGCTTTTACGAAGCAGATGCGAGGAAGATGTTCTGACATA 1020
 QY 1021 CTGCTGCAGGCTCAGTCCAGCTCATTCACAGCAGCAGGTCGCAATGCGAGAGTGGTAC 1080
 DB 1021 CTGCTGCAGGCTCAGTCCAGCTCATTCACAGCAGCAGGTCGCAATGCGAGAGTGGTAC 1080
 QY 1081 CAGGGGGAAGTCAACGAGAAGATGATGTGTGAGGCAATCCCGAAGGGGTGTGGACACC 1140
 DB 1081 CAGGGGGAAGTCAACGAGAAGATGATGTGTGAGGCAATCCCGAAGGGGTGTGGACACC 1140
 QY 1141 TGCCAGGGTCAACAGTGGTGGGCCCTGTATGTACCAATCTGACCAAGTGGCATGTGTGGGC 1200
 DB 1141 TGCCAGGGTCAACAGTGGTGGGCCCTGTATGTACCAATCTGACCAAGTGGCATGTGTGGGC 1200
 QY 1201 ATCTTAGCTGGGGCTATGCTGGGGGGCCCGAGCAGCAGCAGGAGTATACACCAAGGTC 1260
 DB 1201 ATCTTAGCTGGGGCTATGCTGGGGGGCCCGAGCAGCAGGAGTATACACCAAGGTC 1260
 QY 1261 TCAGCCTATCTCAACTGAGTCTCAATGCTGGAAGGTGAGCTG 1305
 DB 1261 TCAGCCTATCTCAACTGAGTCTCAATGCTGGAAGGTGAGCTG 1305

RESULT 2

AR142620
 LOCUS AR142620 2038 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 18 from patent US 6203979.
 ACCESSION AR142620
 VERSION AR142620.1 GI:15103906
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 2038)
 AUTHORS Bandman, O., Hillman, J.L., Yue, H., Guegler, K.J., Corley, N.C.,
 Tang, Y. Tom. and Shah, P.
 TITLE Human protease molecules
 JOURNAL Patent: US 6203979-A 18 20-MAR-2001;
 FEATURES
 source
 1. .2038
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 99.9%; Score 1303.4; DB 6; Length 2038;
 Best Local Similarity 99.9%; Pred. No. 1.8e-301;
 Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGATCTCTGACAGTATCAACCTCTGAAAGCTGTGATCTCAACCCCTGGCGAACCC 60
 DB 200 ATGATCTCTGACAGTATCAACCTCTGAAAGCTGTGATCTCAACCCCTGGCGAACCC 259
 QY 61 CGPATCCCCATGGAGCTTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGC 120

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 08:38:47.; Search time 23 Seconds
(without alignments)
976.404 Million cell updates/sec

Title: US-10-030-688-2
Perfect score: 2342
Sequence: 1 MPDSDQLNSLDVKKLRKP.....VTKVSAIYNWYNWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pdp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pdp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pdp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID
1	2338	99.8	435	US-09-008-271A-6
2	2337	99.8	437	US-09-0851-588-8
3	2186	93.8	423	US-09-656-002-2
4	2188	93.4	406	US-09-851-588-6
5	677.5	28.9	492	US-09-685-166A-895
6	676.5	28.9	492	US-09-342-749-2
7	676.5	28.9	492	US-09-691-840-2
8	655.5	28.0	454	US-09-518-046-2
9	588.5	25.1	417	US-09-820-002-4
10	584	24.9	455	US-09-820-002-2
11	580	24.8	376	US-09-820-002-2
12	574	24.5	416	US-09-820-002-2
13	571	24.4	798	US-09-200-900A-2
14	571	24.4	798	PCT-US94-00616-2
15	568	24.3	418	US-08-508-448C-25
16	568	24.3	418	US-09-370-838-82
17	568	24.3	418	US-09-370-838-82
18	566	24.2	418	US-09-370-838-82
19	558.5	23.8	283	US-08-807-151-1
20	558.5	23.8	283	US-09-478-957-1
21	531.5	22.7	232	US-08-508-448C-19
22	512.5	21.9	256	US-09-027-337-3
23	512.5	21.9	256	US-09-644-600-3
24	512.5	21.9	256	US-09-654-600A-3
25	511	21.8	638	US-08-681-151-3
26	507.5	21.7	255	US-08-944-483-67
27	503.5	21.5	285	US-09-023-942A-26

Sequence 63, Appl
Sequence 6, Appl
Sequence 7, Appl
Sequence 3, Appl
Sequence 257, App
Sequence 257, App
Sequence 257, App
Sequence 4, Appl
Sequence 7, Appl
Sequence 253, App
Sequence 263, App
Sequence 10, Appl
Sequence 7, Appl
Sequence 2, Appl
Sequence 68, Appl
US-08-944-483-63
US-09-023-942A-6
US-09-387-375-7
US-09-008-271A-3
US-09-907-794A-257
US-09-905-125A-257
US-09-902-775A-257
US-09-023-942A-4
US-09-724-675-4
US-09-386-623-7
US-09-907-794A-253
US-09-905-125A-253
US-09-902-775A-253
US-09-644-600-10
US-09-654-600A-10
US-09-386-653A-7
US-09-734-675-2
US-08-944-483-68

ALIGNMENTS

RESULT 1
US-09-008-271A-6
; Sequence 6, Application US/09008271A
; Patent No. 6203979
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tcm Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271A
FILING DATE: 16-Jan-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNNOT13
CLONE: 1337018
SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
US-09-008-271A-6

Query Match 99.8%; Score 2338; DB 3; Length 435;
 Best Local Similarity 99.8%; Pred. No. 1.2e-234;
 Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPSDQPLNSLDVFKPRIPMETFRKVGIPITIIALLSLASIIIVVLKIVLDKYF 60
 DB 1 MDPSDQPLNSLDVFKPRIPMETFRKVGIPITIIALLSLASIIIVVLKIVLDKYF 60

QY 61 LCQOPLHFIIPRKQCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKORSTLOVLD SATGN 120
 DB 61 LCQOPLHFIIPRKQCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKORSTLOVLD SATGN 120

QY 121 WFSACFDNTEALAEACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGP 180
 DB 121 WFSACFDNTEALAEACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGP 180

QY 181 CLSGSLVSLHCLACGKSLKTRPVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLTA 240
 DB 181 CLSGSLVSLHCLACGKSLKTRPVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLTA 240

QY 241 AHCFRKHTDVFNKVRAGSDKLGSPFLSAVAKIIIEFNPMYPKNDIALMKLOPFLTF 300
 DB 241 AHCFRKHTDVFNKVRAGSDKLGSPFLSAVAKIIIEFNPMYPKNDIALMKLOPFLTF 300

QY 301 SGTVRPICLPFDEELTPATLWIIGWFTKQNGGKXSDILLQASVQVIDSTRCNADDAY 360
 DB 301 SGTVRPICLPFDEELTPATLWIIGWFTKQNGGKXSDILLQASVQVIDSTRCNADDAY 360

QY 361 QGVTEKXWCAGIPEGVDTCQDGGGGLMYQSDQHWVGVISWGVGGGSPSTPGVYTKV 420
 DB 361 QGVTEKXWCAGIPEGVDTCQDGGGGLMYQSDQHWVGVISWGVGGGSPSTPGVYTKV 420

QY 421 SAYLNIYNNVKAEL 435
 DB 421 SAYLNIYNNVKAEL 435

RESULT 2
 US-09-851-588-8
 ; Sequence 8, Application US/09851588
 ; Patent No. 6682890
 ; GENERAL INFORMATION:
 ; APPLICANT: Mack, David
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Wilson, Keith E.
 ; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
 ; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
 ; CURRENT APPLICATION NUMBER: US/09/851.588
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: US 09/642,252
 ; PRIOR FILING DATE: 2000-08-17
 ; PRIOR APPLICATION NUMBER: US 09/656,002
 ; PRIOR FILING DATE: 2000-09-06
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 437
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-851-588-8

Query Match 99.8%; Score 2337; DB 4; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.5e-234;
 Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVFKPRIPMETFRKVGIPITIIALLSLASIIIVVLKIVLDKYF 61
 DB 4 DPDSQPLNSLDVFKPRIPMETFRKVGIPITIIALLSLASIIIVVLKIVLDKYF 63

QY 62 CQOPLHFIIPRKQCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKORSTLOVLD SATGN 121

DB 64 CQOPLHFIIPRKQCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKORSTLOVLD SATGN 123
 QY 122 FSAFCDFNTEALAEACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGP 181
 DB 124 FSAFCDFNTEALAEACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGP 183

QY 182 LSGSLVSLHCLACGKSLKTRPVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLTA 241
 DB 184 LSGSLVSLHCLACGKSLKTRPVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLTA 243

QY 242 AHCFRKHTDVFNKVRAGSDKLGSPFLSAVAKIIIEFNPMYPKNDIALMKLOPFLTF 301
 DB 244 AHCFRKHTDVFNKVRAGSDKLGSPFLSAVAKIIIEFNPMYPKNDIALMKLOPFLTF 303

QY 302 GTVRPICLPFDEELTPATLWIIGWFTKQNGGKXSDILLQASVQVIDSTRCNADDAY 361
 DB 304 GTVRPICLPFDEELTPATLWIIGWFTKQNGGKXSDILLQASVQVIDSTRCNADDAY 363

QY 362 GEVTEKXWCAGIPEGVDTCQDGGGGLMYQSDQHWVGVISWGVGGGSPSTPGVYTKV 421
 DB 364 GEVTEKXWCAGIPEGVDTCQDGGGGLMYQSDQHWVGVISWGVGGGSPSTPGVYTKV 423

QY 422 AYLNIYNNVKAEL 435
 DB 424 AYLNIYNNVKAEL 437

RESULT 3
 US-09-656-002-2
 ; Sequence 2, Application US/09656002
 ; Patent No. 6455668
 ; GENERAL INFORMATION:
 ; APPLICANT: Mack, David
 ; APPLICANT: Gish, Kurt
 ; APPLICANT: Wilson, Keith
 ; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
 ; FILE REFERENCE: A-69108/DJB/JJD/AMS
 ; CURRENT APPLICATION NUMBER: US/09/656,002
 ; CURRENT FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: US 09/525,993
 ; PRIOR FILING DATE: 2000-03-15
 ; PRIOR APPLICATION NUMBER: US 09/493,444
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: PCT/US 00/07044
 ; PRIOR FILING DATE: 2000-03-15
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-656-002-2

Query Match 93.8%; Score 2196; DB 4; Length 423;
 Best Local Similarity 97.1%; Pred. No. 6.9e-220;
 Matches 408; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 16 PLRKPRIIPMETFRKVGIPITIIALLSLASIIIVVLKIVLDKYFVLCQOPLHFIIPRKQ 75
 DB 4 PCANPVSEFMRPESVGIPTIALLSLASIIIVVLKIVLDKYFVLCQOPLHFIIPRKQ 63

QY 76 DGEIDCPLGDEEHCVKSPFEGPAVAVRLSKORSTLOVLD SATGNWFSACFDNTEALAE 135
 DB 64 DGEIDCPLGDEEHCVKSPFEGPAVAVRLSKORSTLOVLD SATGNWFSACFDNTEALAE 123

QY 136 TACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGPCLSGSLVSLHCLACG 195
 DB 124 TACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGPCLSGSLVSLHCLACG 183

QY 196 KSLKTRPVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLTAACHCTRKHDTDFNWK 255
 DB 184 KSLKTRPVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLTAACHCTRKHDTDFNWK 243